

ON THE STATISTICAL PROPERTIES OF COSPECTRA

D. HUPPENKOTHEN^{1,2} AND M. BACHETTI³*Draft version September 29, 2017*

ABSTRACT

In recent years, the cross spectrum has received considerable attention as a means of characterising the variability of astronomical sources as a function of wavelength. While much has been written about the statistics of time and phase lags, the cospectrum has only recently been understood as means of mitigating instrumental effects dependent on temporal frequency in astronomical detectors, as well as a method of characterizing the coherent variability in two wavelength ranges on different time scales. In this paper, we lay out the statistical foundations of the cospectrum, starting with the simplest case of detecting a periodic signal in the presence of white noise. This case is especially relevant for detecting faint X-ray pulsars in detectors heavily affected by instrumental effects, including *NuSTAR*, *Astrosat* and *IXPE*. We show that the statistical distributions of both single and averaged cospectra differ considerably from those for standard power spectra. While a single cospectrum follows a Laplace distribution exactly, averaged cospectra are approximated by a Gaussian distribution only for more than ~ 30 averaged segments, dependent on the number of trials. We provide an instructive example of a quasi-periodic oscillation in *NuSTAR* and show that applying standard power spectral statistics leads to underestimated tail probabilities for period detection.

1. INTRODUCTION

Time series analysis is one of the primary ways to understand the physical properties of astronomical object in our universe, from exoplanets and stars to black holes and Active Galactic Nuclei (AGN). Fourier analysis, especially through the periodogram⁴, has long been used to find periodic and quasi-periodic signals as well as characterize the stationary stochastic processes often present in accreting systems. While in principle, the statistics of the periodogram is well understood and characterized in the literature (e.g. van der Klis 1989), the periodogram is often subject to instrumental effects like dead time that change its statistical properties and thus make statistical inference difficult in practice (e.g. Zhang et al. 1995).

In the past years, the field of *spectral timing* has enjoyed significant success by making it possible to combine both temporal and spectral information in a single model. Within this framework, the complex cross spectrum—defined as the Fourier transform of one time series with the complex conjugate of the Fourier transform of a second time series—holds a central position. The cross spectrum is commonly used to compute phase lags, which are easily converted to time lags and are central for understanding e.g. reverberation mapping in accreting black holes (see Uttley et al. 2014, for a recent review).

The real part of the cross spectrum, also named *cospectrum*, has gained less attention, but can be just as useful. It has long been used to study gravity waves in the Earth's

atmosphere (e.g. John & Kishore Kumar 2016), models of the Martian atmosphere (e.g. Wang & Toigo 2016), the solar heliosphere (e.g. Vigeesh et al. 2017), surface elevation of arctic sea ice (e.g. Ardhuin et al. 2016) and drifting snow (e.g. Paterna et al. 2016) as well as surface gravity waves in beaches (e.g. Fiedler et al. 2015) and eddy heat flux in the Earth's troposphere (e.g. Wang & Nakamura 2015; Zurita-Gotor 2017). Within astronomy, it has recently been recognized as one approach to mitigate instrumental effects by making use of the independent detectors present in some telescopes. Where in standard analysis, light curves of multiple detectors are summed before Fourier transforming the summed light curve, it is possible to instead Fourier-transform the signal of two independent detectors. The resulting powers will be less affected by an instrumental effect called dead time (see details in Bachetti et al. 2015), which leads to frequency-dependent changes in the mean and variance of the statistical distributions governing the periodogram. This is relevant for current X-ray detectors such as *NuSTAR* and *Astrosat*, but will also be relevant to future missions with multiple detectors like *IXPE*. This approach has recently been used in *NuSTAR* studies of millisecond pulsars (Ferrigno et al. 2017), Ultraluminous X-ray Sources (Bachetti 2016) and X-ray binaries (Barrière et al. 2015; Zoghbi et al. 2016; Ingram et al. 2016; Huppenkothen et al. 2017; Stiele & Kong 2017). Similarly, the cospectrum of two time series taken with the same instrument, but in different wavelength ranges, can be used to characterize the coherent variability in both time series as a function of frequency.

While much has been written on the subject of the statistics of cross spectra and time lags (e.g. Epitropakis & Papadakis 2016), the derivation of cospectral statistics is notably absent from the astronomy literature, and most publications assume that either the χ^2_2 distribution used for standard periodograms or a Gaussian distribution for averaged spectra is appropriate. That these assumptions are valid has not been shown until now.

¹ Center for Data Science, New York University, 65 5h Avenue, 7th Floor, New York, NY 10003

² Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA

³ INAF-Osservatorio Astronomico di Cagliari, via della Scienza 5, I-09047 Selargius (CA), Italy

⁴ We distinguish in this paper between the *power spectrum*, which describes the process at the source generating variable time series, and the *periodogram*, which denotes a realization of said power spectrum, i.e. the time series we actually observe.

In this paper, we lay out the basic statistical distributions for detecting periodic and narrow quasi-periodic signals in the presence of detector noise (e.g. photon counting statistics) for both single and averaged cospectra. We show below that unlike for the periodogram, the statistical distribution for a single cospectrum reduces to a Laplace distribution, while the distribution for averaged cospectra is considerably more complex. We also find that for averaged cospectra consisting of more than ~ 30 averaged individual segments, the assumption of a Gaussian distribution is indeed appropriate for single-trial tail probabilities and reduces computation overhead. However, for averaged cospectra of fewer segments, more stringent significance thresholds or large numbers of trials, the statistical distribution—and hence the derived p -values used for period detection—deviate significantly from a Gaussian distribution.

The paper is laid out as follows. In Section 2.1, we derive the PDF of a single cospectrum, and show associated simulations and detection thresholds for period detection in Section 2.1.1. Section 2.2 extends the derivation to the common case where the cospectra of several time series are averaged. We show that the statistical distribution indeed changes from a Laplace distribution once multiple cospectra are averaged and becomes consecutively more Gaussian as a larger number of cospectra are included in the average in Section 2.2.1. Finally, Section 3 presents a real-world example using simulated *NuSTAR* data of a quasi-periodic oscillation (QPO) as commonly found in accreting neutron star X-ray binaries. We end in Section 4 with a short discussion and conclusion. All figures and results are reproducible, and the associated code can be found online⁵. In a second, forthcoming paper, we will treat the considerably more complex case of cospectra where the time series consist of stochastic variability and show how to model the cospectrum in both a Maximum Likelihood and Bayesian framework.

For the reader looking for the statistical distributions of relevance, who may be only casually interested in the mathematical background, we point to Equations 15 and 18 for the probability density functions (PDFs) for a single cospectrum and averaged cospectrum, respectively, and Equations 16 and 19 for the cumulative distribution functions (CDFs) in both cases.

2. THE STATISTICAL DISTRIBUTIONS OF COSPECTRAL DENSITIES

2.1. Statistical Distribution for a Single Cospectrum

Consider two independently distributed, constant, stationary time series $\mathbf{x} = \{x_k\}_{k=1}^N$ and $\mathbf{y} = \{y_k\}_{k=1}^N$ with N data points taken at simultaneous time intervals $\{t_k\}_{k=1}^N$. Assume for simplicity that the measurements x_k and y_k are normally distributed, such that $x_k \sim \mathcal{N}(\bar{x}, w_x^2)$ and $y_k \sim \mathcal{N}(\bar{y}, w_y^2)$ with means \bar{x}, \bar{y} and variances w_x^2, w_y^2 . The data points in the time series \mathbf{x} and \mathbf{y} can be expressed in terms of a Fourier series,

$$\begin{aligned} x_k &= \frac{1}{N} \sum_j \mathcal{F}_x(j) \\ y_k &= \frac{1}{N} \sum_j \mathcal{F}_y(j) \end{aligned} \quad (1)$$

where

$$\mathcal{F}_x(j) = \frac{1}{2}(A_{xj} - iB_{xj})e^{-i(\frac{2\pi jt}{T})} \quad (2)$$

$$\mathcal{F}_y(j) = \frac{1}{2}(A_{yj} - iB_{yj})e^{-i(\frac{2\pi jt}{T})}, \quad (3)$$

where $i = \sqrt{-1}$, and A_{xj}, A_{yj} and B_{xj}, B_{yj} describe the real and imaginary parts of the Fourier amplitudes, respectively (for a pedagogical introduction into Fourier analysis, see van der Klis 1989). The complex cross spectrum is then calculated by multiplying the Fourier transform of light curve \mathbf{x} with the complex conjugate of the Fourier transform of light curve \mathbf{y} (Vaughan & Nowak 1997; Nowak et al. 1999, see also Uttley et al. 2014 for a recent review of spectral timing techniques):

$$\begin{aligned} \mathcal{F}_x(j)\mathcal{F}_y^*(j) &= \frac{1}{2}(A_{xj} - iB_{xj})e^{i\frac{2\pi jt}{T}} \frac{1}{2}(A_{yj} + iB_{yj})e^{i\frac{-2\pi jt}{T}} \\ &= \frac{1}{4}[(A_{xj}A_{yj} + B_{xj}B_{yj}) + \\ &\quad i(A_{xj}B_{yj} - A_{yj}B_{xj})] \end{aligned} \quad (4)$$

Note that for strictly real-valued time series, as light curves in astronomy always are, $A_j = A_{-j}$ and $B_j = -B_{-j}$, such that

$$\mathcal{F}_x(j)\mathcal{F}_y^*(j) = \frac{1}{2} \{ (A_{xj}A_{yj} + B_{xj}B_{yj}) + i(A_{xj}B_{yj} - A_{yj}B_{xj}) \}. \quad (5)$$

The real part of this equation is the cross-spectral equivalent of the power spectral density, also called the cospectrum:

$$C_j = \frac{1}{2}(A_{xj}A_{yj} + B_{xj}B_{yj}). \quad (6)$$

For normally distributed light curves (and indeed for uncertainties coming from a wide range of statistical distributions, including the Poisson distribution, if N is large), the real and imaginary amplitude components are distributed as $A_{xj}, B_{xj} \sim \mathcal{N}(0, \sigma_x^2)$ with $\sigma_x = \sqrt{\frac{\sum_{k=1}^N x_k^2}{2}}$ and $A_{yj}, B_{yj} \sim \mathcal{N}(0, \sigma_y^2)$ with $\sigma_y = \sqrt{\frac{\sum_{k=1}^N y_k^2}{2}}$. For standard periodograms, $A_{xj} = A_{yj}$ and $B_{xj} = B_{yj}$, and the power spectral density reduces to $P_j = \frac{1}{2}(A_j^2 + B_j^2)$, which is well-known to follow a χ^2 distribution with 2 degrees of freedom.

Because this condition is not fulfilled for cospectra, we need to derive the probability distribution of the sum of the products over *independent* Gaussian random variables. The probability distribution of the product of two random

⁵ <https://github.com/dhuppenkothen/cospectra-paper>

variables⁶ $Z = A_{xj}A_{yj}$ is called the product distribution, defined as

$$p_Z(z) = \int_{-\infty}^{+\infty} p_X(x) p_Y\left(\frac{z}{x}\right) \frac{1}{|x|} dx \quad (7)$$

$$= \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) \exp\left(-\frac{(z/x)^2}{2\sigma_y^2}\right) \frac{1}{|x|} dx.$$

It can be shown (Watson 1922; Wishart & Bartlett 1932) that the integral in Equation 8 above can be reduced to

$$P_Z(z) = \frac{K_0\left(\frac{|z|}{\sigma_x\sigma_y}\right)}{\pi\sigma_x\sigma_y}, \quad (8)$$

where $K_0(x) = \int_0^{+\infty} \frac{\cos(xt)}{\sqrt{t^2+1}} dt$ is the Bessel function of the second kind of order 0. We can now use this result to derive the probability density function of C_j . In particular, we find that both random variables $Z_j = A_{xj}A_{yj}$ and $Q_j = B_{xj}B_{yj}$ follow the Bessel distribution defined in Equation 8. Our task is therefore to find the PDF of the sum of two Bessel distributions. The PDF of this sum requires the convolution of the PDFs of each individual random variable being summed. This convolution is difficult to calculate directly for the Bessel distribution defined in Equation 15 above. We instead consider the *moment-generating function* of the PDF, generally defined as

$$M_Z(t) := \mathbb{E}[e^{tZ}] \quad (9)$$

for a random variable Z . Consider the sum of any two independent random variables, $S = Z + Q$. While the PDF of S can be found via the convolution of the individual PDFs, it is often simpler to consider the moment-generating function, where the convolution reduces to a simple multiplication operation:

$$M_S(t) = M_Z(t)M_Q(t). \quad (10)$$

The moment-generating function of the Bessel distribution in Equation 8 above is, in the general case (Seijas-Macías & Oliveira 2012) where the means μ_x and μ_y are non-zero and the random variables may have unequal variances $\sigma_x \neq \sigma_y$:

$$M_Z(t) = \frac{\exp\left(\frac{t\mu_x\mu_y + 0.5(\mu_y^2\sigma_x^2 + \mu_x^2\sigma_y^2)t}{1 - t^2\sigma_x^2\sigma_y^2}\right)}{\sqrt{1 - t^2\sigma_x^2\sigma_y^2}}, \quad (11)$$

but since $\mu_x = \mu_y = 0$ for Fourier amplitudes, this reduces to

$$M_Z(t) = \frac{1}{\sqrt{1 - t^2\sigma_x^2\sigma_y^2}}, \quad (12)$$

Thus, the moment-generating function for the sum of Z_j and Q_j becomes

$$M_C(t) = M_Z(t)M_Q(t) = \frac{1}{1 - t^2\sigma_x^2\sigma_y^2}. \quad (13)$$

⁶ We continue the following derivation using the amplitudes A_j , but the same arguments apply exactly to the imaginary amplitudes B_j .

We note that the Laplace distribution is defined as

$$p_{\text{Laplace}}(x|\mu, b) = \frac{1}{2b} \exp\left(-\frac{|x - \mu|}{b}\right)$$

and its moment-generating function as

$$M_{\text{Laplace}}(t) = \frac{e^{t\mu}}{1 - b^2 t^2}.$$

Comparing this last equation with Equation 13, we find that that Equation 13 is equal to the moment-generating function of the Laplace distribution with $\mu = 0$ and $b = \sigma_x\sigma_y$, and hence the cospectral densities follow a Laplace distribution:

$$p(C_j|0, \sigma_x\sigma_y) = \frac{1}{\sigma_x\sigma_y} \exp\left(-\frac{|C_j|}{\sigma_x\sigma_y}\right) \quad (14)$$

with

$$\sigma_x = \sqrt{\frac{\sum_{k=1}^N x_k}{2}} \quad \text{and} \quad \sigma_y = \sqrt{\frac{\sum_{k=1}^N y_k}{2}}. \quad (15)$$

2.1.1. Detection Thresholds

Detection thresholds for cospectra will generally be different from those of classical power spectra, because the Laplace distribution tends to be narrower than the equivalent χ^2_2 distribution for single power spectra. To show how the distributions and the corresponding detection thresholds differ, we simulated simple Poisson-distributed light curves. First, we simulated two light curves with a duration of 10 s and 10^6 data points each, corresponding to a time resolution of 10^{-5} s. The light curves have an identical mean count rate of 10^6 counts/s, corresponding to 10 counts per bin. In order to simulate typical measurement uncertainties, we sampled from a Poisson distribution for each time bin with a rate parameter $\lambda = 10$, corresponding to the average counts per bin.

We then calculated both the cospectrum of the two light curves and the periodogram of only the first light curve for comparison. For simplicity, both spectra were computed in Leahy normalization (Leahy et al. 1983), which is typically used when searching for (quasi-)periodic signals in time series. In order to normalize the cospectrum correctly, we used $2/\sqrt{N_{\text{ph},x}N_{\text{ph},y}}$, where $N_{\text{ph},x}$ and $N_{\text{ph},y}$ are the number of photons of light curves \mathbf{x} and \mathbf{y} , respectively, as prescribed by Bachetti et al. (2015). In this normalization the powers are distributed as χ^2_2 exactly for the periodogram, and following a Laplace distribution with $\mu = 0$ and $\sigma = 1$ for the cospectrum. In Figure 1, we plot the resulting distribution of powers. While the periodogram is only defined for positive values, the Laplace distribution is symmetric around zero, and in general the cospectrum will comprise both positive and negative powers. It is also immediately visible from Figure 1 that the probability of obtaining a certain (positive) noise power is always lower for the Laplace distribution than for the χ^2_2 distribution. In practice, this implies that using the latter where the former is appropriate, we may miss significant periodic signals, because we assume them to be weaker than they are in reality. To demonstrate this, we plot the survival function in Figure 2. The survival function,

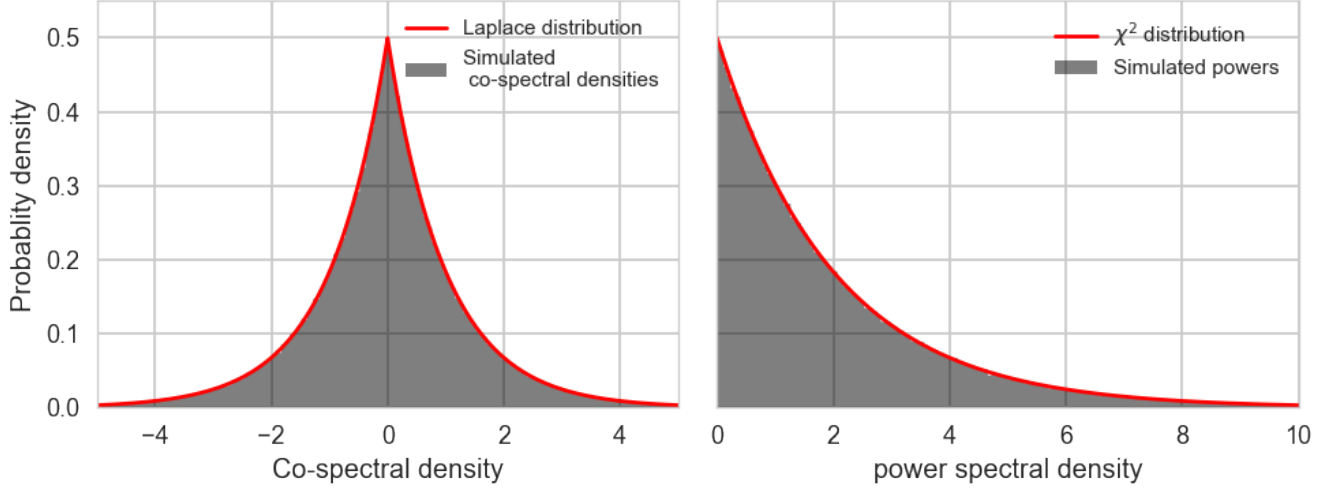


FIG. 1.— Distribution of Leahy-normalized cospectral densities (left) and power spectral densities (right), respectively, for the simulated data. In dark grey, we show fine-grained histograms of the simulated powers. In red we plot the theoretical probability distribution the simulated powers should follow: A Laplace distribution with $\mu = 0$ and $\sigma = 2$ for the cospectral densities and a χ^2 -distribution with 2 degrees of freedom for the power spectral densities. The simulated powers adhere very closely to the theoretical predictions.

defined in terms of the CDF as $SF(x) = 1 - CDF(x)$, encodes the tail probability of seeing at least a value $x \geq X$. This tail probability is often considered to be the p -value of rejecting the null hypothesis that a certain candidate for a periodic signal could be reasonably produced by the noise powers. The CDF for the Laplace distribution with $\mu = 0$ is defined as

$$F_{C_j}(x) = \begin{cases} \frac{1}{2} \exp\left(-\frac{C_j}{\sigma_x \sigma_y}\right) & \text{if } C_j < 0 \\ 1 - \frac{1}{2} \exp\left(-\frac{C_j}{\sigma_x \sigma_y}\right) & \text{if } C_j \geq 0 \end{cases} \quad (16)$$

Much like the PDF, the tail probability is always smaller for the Laplace distribution, indicating that for a given candidate signal, the p -value for rejecting the null hypothesis will be stronger than for χ^2 -distributed variables. To reinforce this statement, we again simulated two light curves, each again with a duration of 10s, but this time with only 1000 data points for simplicity and speed, and a time resolution of 0.01s. For this simulation, we assumed a mean count rate of 1000 counts/s or 10 counts per bin, and additionally introduced a sinusoidal signal with a period of 0.1s and a fractional rms amplitude of $a_{\text{frac}} = 0.055$. Again, this template was used to produce two Poisson-distributed light curves with a rate parameter equal to the number of counts in each bin as defined by the flat continuum and the periodic signal. In Figure 3, we show the cospectral densities along with trial-corrected 0.99 detection thresholds for both the Laplace and χ^2 distribution. If the powers are assumed to follow a χ^2 distribution, as for the periodogram, the candidate at 10Hz would be discounted at the 99% detection threshold, whereas correctly applying the Laplace distribution yields a correct rejection of the null hypothesis at the same detection threshold.

2.2. Averaged Cospectra

The χ^2 distribution used for periodograms has the simple property that sums of χ^2 -distributed variables

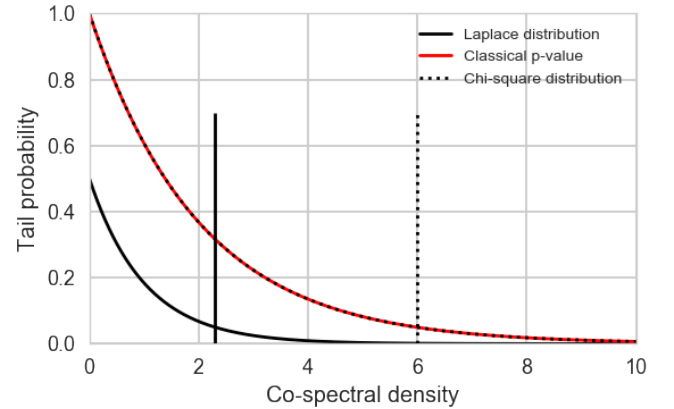


FIG. 2.— Tail probabilities for the Laplace and χ^2 distributions, respectively. The tail probability, or survival function, is defined as $SF(x) = 1 - CDF(x)$. The tail probability measures the probability of observing a value $x \geq X$, and is often used for detecting periods in power spectra. For the power spectral densities, we plot both the theoretical prediction for the survival function based on the χ^2 distribution (black dashed line), as well as the corrected distribution for power spectra derived in Groth (1975) (red solid line). For illustrative purposes, we show a single-trial 95% detection threshold for the Laplace distribution (black solid vertical line) and the χ^2 distribution (black dashed vertical line).

again follow the same distribution, with a different number of degrees of freedom. The same is not true for the Laplace distribution. For n independent and identically distributed (i.i.d.) random variables distributed following a standard Laplace distribution with a mean of $\mu = 0$ and a width of $b = 1$, the distribution of the sums of these random variables can be derived using the fact that a single Laplace random variable X can be rewritten as the difference of two exponential random variables,

$$X = Z - Z',$$

and thus for n summed random Laplace random variables,

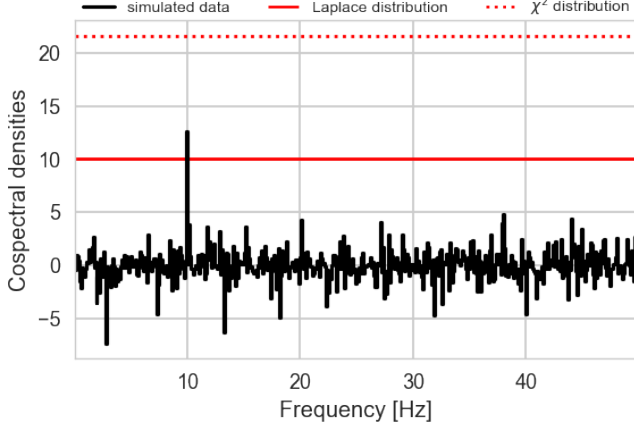


FIG. 3.— Cospectrum of two simulated light curves, each with a constant continuum flux of 10 counts per bin and a periodic signal at 10 Hz. The latter is clearly visible in the cospectrum. We also show the 99% detection threshold, corrected for a number of trials equal to the number of spectral bins, assuming Laplace-distributed data (red solid line) and χ^2 -distributed data (red dashed line). When the latter distribution is assumed, the periodic signal would not be considered a significant detection, because the χ^2 distribution produces a wider distribution of powers. Applying the correct Laplace distribution, however, allows for the detection of weaker signals.

$$T = \sum_{i=1}^n X_i = \sum_{i=1}^n Z_i - \sum_{i=1}^n Z'_i = G_1 - G_2, \quad (17)$$

where G_1 and G_2 are i.i.d. standard gamma random variables with a distribution $g(x) = \frac{1}{\Gamma(\nu)} x^{\nu-1} e^{-x}$ and a shape parameter $\nu = n$. For the full derivation of the density, we refer the reader to Kotz et al. (2001) and simply state the end result for the PDF for n averaged standard Laplace random variables, \bar{X}_n (see Kotz et al. 2001, Equations 2.3.25 and 2.3.26):

$$f_{\bar{X}_n}(x) = \frac{ne^{-|nx|}}{(n-1)!2^n} \sum_{j=0}^{n-1} \frac{(n-1+j)!}{(n-1-j)!j!} \frac{|nx|^{n-1-j}}{2^j}, \quad x \in \mathbb{R}. \quad (18)$$

For practical purposes, evaluating this PDF for averaged spectra above $n \sim 85$ is difficult numerically, because the factorials and exponents in the sum become very large and small, respectively. However, as we will show in Section 2.2.1 below, we find that in practice, when $n \gtrsim 30$, detection thresholds derived from Equation 18 provide only a negligible difference over that derived from a normal distribution $N(0, \sqrt{2/(n+1)})$, depending on the significance threshold required and the number of trials.

In order to derive tail probabilities useful for hypothesis testing, we require the CDF rather than the PDF. In order to correctly account for the absolute values in the PDF, we split the CDF into two parts: a case where $x < 0$ and a case where $x \geq 0$. The integral $F_{\bar{X}_n}(x) = P(X \leq x) = \int_{-\infty}^x f_{\bar{X}_n}(t) dt$ then becomes:

$$F_{\bar{X}_n}(x) = \begin{cases} \sum_{j=0}^{n-1} D \frac{1}{n} (2\Gamma(-j+n) - \gamma(-j+n, nx)), & x \geq 0 \\ \sum_{j=0}^{n-1} D \frac{1}{n} \gamma(-j+n, -nx), & x < 0 \end{cases} \quad (19)$$

where $\Gamma(l) = (l-1)!$ is the gamma function, $\gamma(l+1, m) = l\gamma(l, m) - l^m e^{-m}$ is the incomplete upper gamma function, and the pre-factor constant D is defined as

$$D = \frac{n(n-1+j)!}{(n-1-j)!(n-1)!2^{n+j}}.$$

As laid out in Section 2.1.1, the tail probability can easily be calculated via the survival function, $\text{SF}(x) = 1 - \text{CDF}(x)$.

2.2.1. Detection Thresholds

In order to show the way the probability distribution changes as a function of averaged cospectra, we simulate light curves of 10^5 data points and a mean count rate of 100 counts/s consisting of pure white noise. We compute n such light curves and average their cospectral powers in order to show the distribution of those powers compared to the expected probability distributions. In Figure 4, we show the simulated distribution of Leahy-normalized powers, along with the distributions that describe them. For a single, non-averaged spectrum, we use the Laplace PDF described in Equation 14. When averaging 10 cospectra, we use Equation 18 and show that the theoretical predictions agree with the simulated powers. Finally, for an averaged cospectra consisting of 100 individual light curves, Equation 18 becomes difficult to compute numerically, and we use a Gaussian PDF instead, which describes the distribution of simulated powers well.

In order to assess the effect on the p -values derived from averaged cospectra, we calculate the tail probabilities for the simulated data sets and compare them with the theoretically expected survival function as defined in Equation 19 as well as a simple Gaussian distribution (Figure 5). Similar to the results derived by Balakrishnan & Kocherlakota (1986), we find that for cospectra of more than 30 averaged light curves, a Gaussian distribution yields a reasonably good approximation to the true distribution up to $p \approx 10^{-4}$ with lower overhead. Note, however, that this holds for single-trial probabilities. In general, one will wish to correct for calculating the significance of multiple trial frequencies, requiring the use of more stringent significance threshold. As shown in Figure 5, the tail probabilities diverge as a function of power, and thus the Gaussian approximation will increasingly overestimate the significance of the signal the higher the threshold is set. Depending on the number of trials used, it is hence advisable to use Equation 19 as long as it remains numerically stable (depending on implementation, up to ~ 85 averaged cospectra).

3. EXAMPLE: A QPO IN NuSTAR

In order to show the difference of the detection limits with the cospectrum and the power spectrum, we show how a QPO at 200 Hz from a very bright source would appear in NuSTAR. We simulate a light curve of $T = 3000$ s duration with a time resolution of $\delta t = 0.5$ ms and an average count rate of 200 counts s^{-1} . To this constant background we add a quasi-periodic oscillation with a

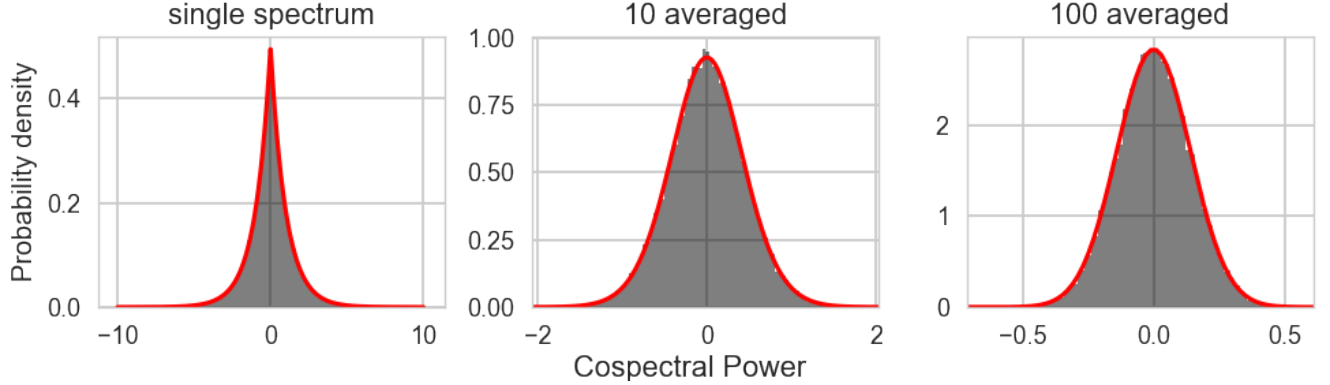


FIG. 4.— Histogram of cospectral powers (grey) and the theoretical expectation (red) for three different cases. In the right panel, we show the distribution of powers for a single cospectrum, with its expected Laplace distribution from Equation 14, in the middle and right panel cospectral powers for averaging 10 (middle) and 100 (right) individual cospectra together. In the middle panel, the theoretical expectation of the sampling distribution is given by Equation 18, in the right panel by a Gaussian distribution with $\sigma = \sqrt{2/101}$.

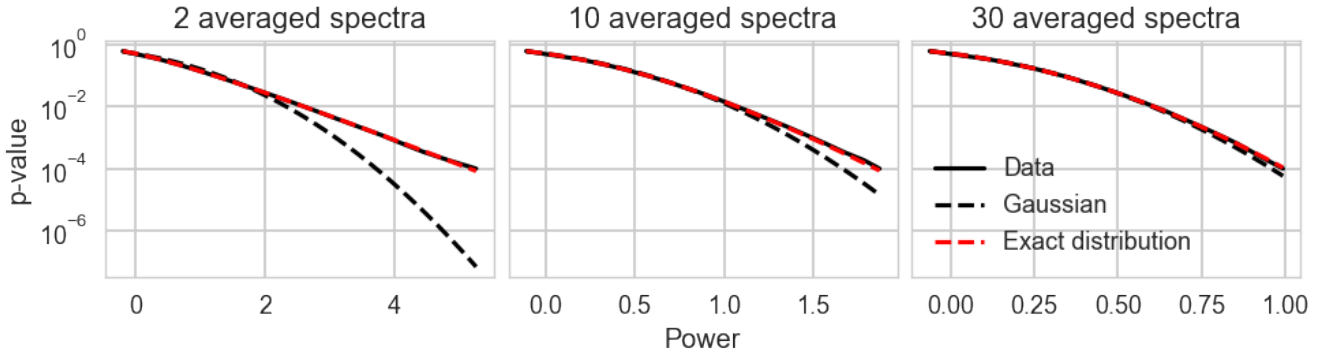


FIG. 5.— Predictions for single-trial p -values as a function of cospectral power, for the simulated data sets (black solid line) compared to the theoretically predicted p -values using the CDF derived in Equation 19 (red dashed line) as well as the survival function of a Gaussian distribution (black dashed line). The exact distribution matches the empirically derived tail probabilities from simulations within the uncertainties of the simulations. For the case of two averaged cospectra (left panel), a Gaussian approximation is obviously the wrong choice, and will lead to vastly overestimated significances, because the Gaussian PDF drops off much more sharply than the exact distribution. However, as more cospectra are averaged, the resulting distribution becomes more and more similar to that of a Gaussian (middle panel), and for ~ 30 averaged cospectra, a Gaussian survival function yields a reasonably good approximation to the tail probabilities one would derive from simulations (right panel), up to $p \approx 10^{-4}$.

period of 5 ms, a fractional rms amplitude of $f_{\text{rms}} = 0.15$ and phases randomized using a normal distribution with a mean of 0 and a width of $\sigma_{\text{qpo}} = 0.01$. We then simulated photon events using rejection sampling from this light curve using the software package *stingray*⁷, running the function `stingray.Eventlist.simulate_times()` twice in order to produce two light curves that are statistically independent, but have the same signal and properties, as we would expect from an instrument with two independent detectors observing the same object. Subsequently we simulated variable deadtime for both light curves with an average time scale of 2.5 ms as commonly seen in *NuSTAR* data (Bachetti et al. 2015) using *HENDRICS*⁸ (Bachetti 2015). We then produced the periodogram of the summed light curves, the averaged periodogram of the two individual light curves, and the cospectrum of the two light curves. Note that in all three cases we produced

averaged periodograms and co-spectra by splitting the light curves into 600 segments of 5 s length each in order to suppress the variance in the powers and show the effects of dead time more clearly.

The results are shown in Figure 6. While in all three cases, the QPO is clearly visible, the two periodograms show strong deviations from the expected flat power spectrum. The shape is distorted and requires a precise model of the non-linearly increasing baseline with a non-linearly increasing rms. While in principle, the periodogram of the combined light curves would have a higher significance by a factor of $\sqrt{2}$, the modeling requirements complicate the calculation of the significance of the QPO. The baseline of the cospectrum, conversely, is not distorted by dead time, and requires only an estimate of the *local* rms in order to calculate the significance of the QPO using Equation 19: it is sufficient to *multiply the cospectrum around the feature by an estimate of the local standard deviation of the white noise* (which is 1 in the ideal case) to use the equations above with no modifications.

⁷ <https://github.com/StingraySoftware/stingray>

⁸ <https://github.com/StingraySoftware/HENDRICS>

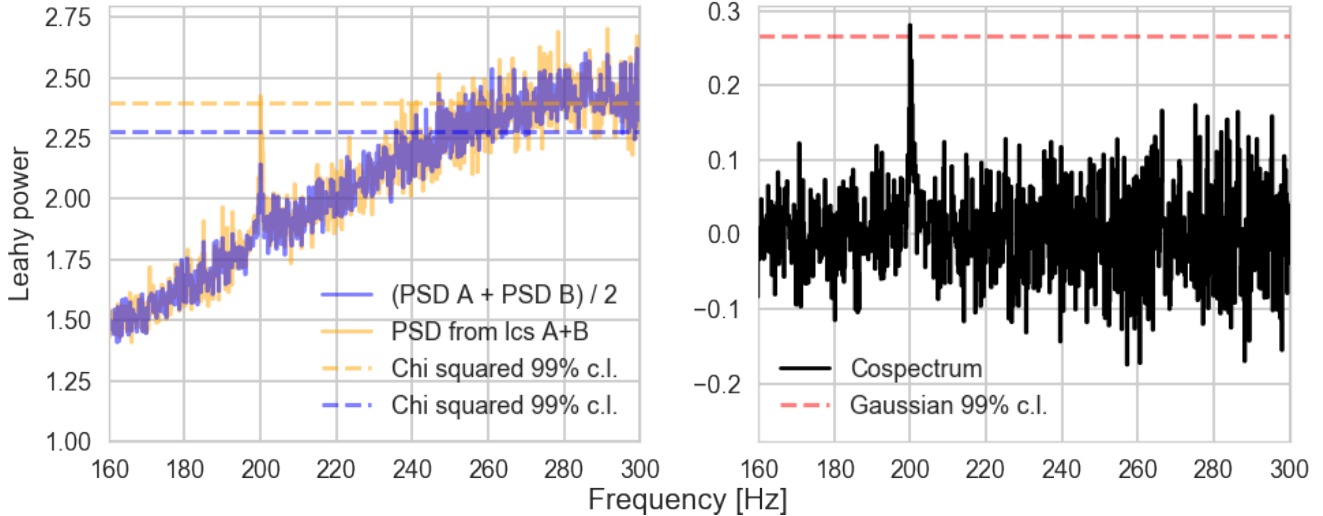


FIG. 6.— Comparison of the averaged periodograms from the two detectors (blue), the periodogram obtained by the sum of the light curves (yellow; both left panel) and the cospectrum (right panel), for 3000 s of synthetic *NuSTAR* data with an incident count rate of 200 count s^{-1} and a strong QPO at 200 Hz. The QPO has an rms of $\sim 9\%$ and a high Q factor of ~ 40 . The shape of the periodogram is strongly distorted from the expected flat powerspectrum centred on a constant value of 2, with deviations of more than 0.5 in the mean level of the powers common. In dashed lines, we also show the expected (trial-corrected) 99% confidence level for the periodograms (in the same colour, left panel) as well as the cospectrum (red, right panel). The signal is detected in both the periodogram of the combined light curves as well as the cospectrum. However, because the white noise level is variable in the periodogram and the power spectrum dips below the expected level of 2 around the frequency of the QPO, the signal is not detected as significantly as it should have been if no dead time was present. Assessing the significance is much more straightforward in the cospectrum, which retains a flat baseline.

4. DISCUSSION AND CONCLUSIONS

We have derived the statistical distribution for the cospectrum, defined as the real part of the cross spectrum. We show that because the Fourier amplitudes being multiplied to derive the cospectrum are now no longer identical (as is the case in the periodogram), the statistical distributions no longer reduce to a simple χ^2 distribution with two degrees of freedom. Instead, we find that the powers in a single cospectrum follow a Laplace distribution with a mean of $\mu = 0$ and a width of $\sigma = 1$. This has important consequences for period detection. Most importantly, the Laplace distribution is considerably narrower than the χ^2_2 distribution expected for periodograms, and thus the significance of a candidate periodic signal will generally be underestimated when using the latter. Using the correct distribution therefore helps correctly identifying weak signals, which the χ^2_2 distribution would ignore as false negatives. Similarly, we find that the sums of Laplace distributions do not follow a similarly simple expression as in the periodogram case, but is considerably more computationally expensive, and may be difficult to estimate numerically when the number n of averaged light curves in the final cospectrum is large. However, we find that for $n \gtrsim 30$, the statistical distribution can be well approximated by a Gaussian distribution, and the resulting tail probabilities used for period detection are very nearly the same as those derived from the exact distribution, up to a tail probability of $p \approx 10^{-4}$. This conclusion, however, depends sensitively on the detection threshold as well as the number of trials: for very small tail probabilities, the two distributions may still deviate significantly. For practical purposes, we suggest using Equation 19 for at least up to ~ 30 averaged cospectra, but also for averages of more spectra if significance thresholds for tail probabilities are smaller than 10^{-4} , or the

number of trials is large.

As an example, we have simulated how a QPO would appear in *NuSTAR* in the presence of dead time, and have shown that the shape of the periodogram is strongly distorted, whereas that of the cospectrum is not (for a longer introduction into the cospectrum and how it can be used in the presence of dead time, see also Bachetti et al. 2015). The significance of the QPO is difficult to assess in the periodogram, because of the non-linearities introduced into the power spectrum by the variable dead time. The standard χ^2_2 distributions may either overestimate or underestimate the significance, depending on the shape of the power spectrum at a given frequency, adding complexity to the detection process in the form of finding a non-linear model for the dead time. The cospectrum, on the other hand, only requires an estimate of the local variance in order to use the equations derived above, making it a far more convenient choice for periodicity detection.

The distributions laid out above allow for detecting periodic and narrow quasi-periodic signals in the presence of detector white noise, and especially important in the context of pulsar detection in X-rays, where faint sources may yield marginal detections even in the best of cases. At the same time, as instruments like *Astrosat* and *IXPE* allow for observations with higher sensitivity, incorporating an accurate treatment of instrumental biases becomes increasingly important, and the cospectral statistics laid out here provide powerful tools to do so. Notably absent from this discussion, however, is the much more common case where a source exhibits stochastic variability in the form of red noise or notably broadened quasi-periodic oscillations. In this case, the goal is either estimation of the precise properties of the underlying stochastic process, or detection of periods against a background varying

stochastically. As has been shown above, the fact that the cross spectrum consists of two different time series complicates the statistical distributions considerably, and this is similarly true cospectra with variability. The exact treatment of this case is beyond the scope of this paper, and will be considered in depth in a forthcoming publication.

Acknowledgements— The authors thank Thomas Laetsch for helpful suggestions regarding the mathematical deriva-

tions. Daniela Huppenkothen is supported by the James Arthur Postdoctoral Fellowship and the Moore-Sloan Data Science Environment at New York University. MB is supported in part by the Italian Space Agency through agreement ASI-INAF n.2017-12-H.0 and ASI-INFN agreement n.2017-13-H.0.

REFERENCES

- Ardhuin, F., Sutherland, P., Doble, M., & Wadhams, P. 2016, *Geophys. Res. Lett.*, 43, 5775
- Bachetti, M. 2015, *MaLTPyNT: Quick look timing analysis for NuSTAR data*, Astrophysics Source Code Library, ascl:1502.021
- . 2016, *Astronomische Nachrichten*, 337, 349
- Bachetti, M., Harrison, F. A., Cook, R., et al. 2015, *ApJ*, 800, 109
- Balakrishnan, N., & Kocherlakota, S. 1986, *Sankhyā: The Indian Journal of Statistics, Series B (1960-2002)*, 48, 439
- Barrière, N. M., Krivonos, R., Tomsick, J. A., et al. 2015, *ApJ*, 799, 123
- Epitropakis, A., & Papadakis, I. E. 2016, *A&A*, 591, A113
- Ferrigno, C., Bozzo, E., Sanna, A., et al. 2017, *MNRAS*, 466, 3450
- Fiedler, J. W., Brodie, K. L., McNinch, J. E., & Guza, R. T. 2015, *Geophys. Res. Lett.*, 42, 9933
- Groth, E. J. 1975, *ApJS*, 29, 285
- Huppenkothen, D., Younes, G., Ingram, A., et al. 2017, *ApJ*, 834, 90
- Ingram, A., van der Klis, M., Middleton, M., et al. 2016, *MNRAS*, 461, 1967
- John, S. R., & Kishore Kumar, K. 2016, *Journal of Atmospheric and Solar-Terrestrial Physics*, 138, 74
- Kotz, S., Kozubowski, T. J., & Podgórski, K. 2001, *Asymmetric Multivariate Laplace Distribution* (Boston, MA: Birkhäuser Boston), 239–272
- Leahy, D. A., Darbro, W., Elsner, R. F., et al. 1983, *ApJ*, 266, 160
- Nowak, M. A., Vaughan, B. A., Wilms, J., Dove, J. B., & Begelman, M. C. 1999, *ApJ*, 510, 874
- Paterna, E., Crivelli, P., & Lehning, M. 2016, *Geophys. Res. Lett.*, 43, 4441
- Seijas-Macías, A., & Oliveira, A. 2012, *Discussiones Mathematicae Probability and Statistics*, 32, 87
- Stiele, H., & Kong, A. K. H. 2017, *ApJ*, 844, 8
- Uttley, P., Cackett, E. M., Fabian, A. C., Kara, E., & Wilkins, D. R. 2014, *A&A Rev.*, 22, 72
- van der Klis, M. 1989, in *NATO Advanced Science Institutes (ASI) Series C, Vol. 262, NATO Advanced Science Institutes (ASI) Series C*, ed. H. Ögelman & E. P. J. van den Heuvel, 27
- Vaughan, B. A., & Nowak, M. A. 1997, *ApJ*, 474, L43
- Vigeesh, G., Jackiewicz, J., & Steiner, O. 2017, *ApJ*, 835, 148
- Wang, H., & Toigo, A. D. 2016, *Icarus*, 271, 207
- Wang, L., & Nakamura, N. 2015, *Geophys. Res. Lett.*, 42, 8192
- Watson, G. N. 1922, *A treatise on the theory of Bessel functions* (Cambridge, University Press)
- Wishart, J., & Bartlett, M. S. 1932, *Mathematical Proceedings of the Cambridge Philosophical Society*, 28, 455?459
- Zhang, W., Jahoda, K., Swank, J. H., Morgan, E. H., & Giles, A. B. 1995, *ApJ*, 449, 930
- Zoghbi, A., Miller, J. M., King, A. L., et al. 2016, *ApJ*, 833, 165
- Zurita-Gotor, P. 2017, *Geophys. Res. Lett.*, 44, 2007